**Variational Principle**

The typical RSPT, or BWPT approaches gives good results when the perturbation is small. If the perturbation isn’t small, then we can use the WKB approximation, but only when we’re in the semi-classical region, i.e., in the limit of high quantum numbers (though sometimes it even gives exact results for low quantum numbers). What if the perturbation isn’t small, and we’re not in the semi-classical region? Then we have recourse to the Variational Principle, which happens to work best in the limit of low quantum numbers. Its employ is close in spirit to the variational least action principle which gives trajectory of a particle in classical mechanics. So we will likewise formulate a variational principle for the wavefunctions. Consider the functional which gives the expectation of H, constrained so that the wave functions are normalized to unity. If we were to determine the functions which do this then we’d calculate,



Taking the functional derivative w/r to ψ\* and setting to 0 so as to extremize H we have,



which is the Schrödinger equation.



So the solutions to the Schrodinger equation are functions which extremize the functional, F, and eigenvalues are just the Lagrange multipliers that enforce the normalizability constraint. So a method to approximate the ψ’s, and λ’s is to minimize the expectation of H. So to get a good approximation of the ground state, we parameterize a normalized function and minimize the energy expectation. We can extend the method to find excited states. To get the 1st excited state we would parameterize a wavefunction orthogonal to our ground state, and then minimize F with it. To get the 2nd excited state we’d parameterize a wavefunction orthogonal to our ground state and 1st excited state, and again minimize F with it. We can repeat similarly to find higher excitations, but as one might expect, the errors would build cumulatively as we go to higher and higher excited states. So because of this, the Variational Principle is best for low lying states.

I surmise there is another way to formulate the variational principle as a typical eigenvalue problem. Say we think the ground state and low lying excited state are well approximated by some linear combination of states |j>. So then we’d have an ansatz for these states:



where |j> are the known wavefunctions, which are linearly independent, though not necessarily orthogonal, and aj are coefficients to be determined. To work out the estimate for the GS we’d normalize our |ψ> and calculate the energy expectation <ψ|H|ψ>, and minimize it w/r to the unknown parameters aj. Then to get say the first excited state, if we wanted to use this same vector space of wavefunctions, we subtract off the projection of the ground state onto |ψ>, and redo the whole procedure with our modified ansatz |ψ´> = |ψ> - |GS><GS|ψ> But I think an equivalent, faster way to do all that is to solve the Schrodinger equation projected onto the vector space of wavefunctions that |ψ> comprises. In other words, we’d do:



This would be our matrix equation for the aj’s. Once solved, we will get M eigenvalues and M states which would be the approximate M lowest lying energies and states of our H. And by approximate, I mean, it would (I think) coincide with what we’d get via use of the variational principle. Certainly, if we allowed our ansatz conjecture to cover the whole Hilbert Space, and not just a subspace of M states, then it must be the case that our result using this procedure would be exactly correct, and it must also be the case that our result using the Variational Principle would be exactly correct. So these two procedures certainly match up in that extreme case. Parenthetical note: observe in the last line above, we cannot say <i|j> = δij unless the |j>’s are all orthonormal. Often this is not the case, though we could use a Gram-Schmidt process from Linear Algebra to find an orthonormal basis for the vector space spanned by |ψ>.

Often the overlap between the states |j> is small, and is considered negligible. In that case, we’ll just have:



which is an easier equation to work out. This approach is the one taken – not in so many words – when using the 2nd quantized tight binding approximation in Condensed Matter physics.

**Example. Hydrogen atom**

Let’s see if we can approximate the ground state energy of the hydrogenic atom. So,



and,



We know that in general bound state have an exponential decay envelope, so we might postulate:



First we need to normalize it so,



So then,



Now let’s form the expectation of H and minimize it:



Still goin,



Okay. Now need to minimize expectation of H. So taking derivative w/r to a,



Looking back at our exact solution of the Hydrogenic atom, we can recognize this as:



Now let’s go back and fill this into the energy,



Comparing to what we got when we did the Hydrogenic atom exactly,



we see this is indeed the Ground state energy of the hydrogenic atom. Of course this was facilitated by our having guessed the exact form of the ground state wave function for the hydrogenic atom in the beginning. If we had guessed something else, we would’ve found an energy higher, but still, probably pretty close.

**Example**

Consider a Hamiltonian,



Use trial wavefunctions of the form,



to work out the best estimate for the ground state energy. First need the normalization,



So now our wavefunction is:



Minimize the energy expectation of ψ(r),



We’ll note the delicate cancelation of the imaginary parts. And,



So we have:



That’s interesting. So if G is large, then we’d just set kI to ∞. And if G is small, then we’d set kI to 0, and kR too. Hmmm….kind of pathological.

**Example**

Consider an anharmonic oscillator:



We’d like to estimate the ground state wavefunction. We’ll estimate the wavefunction as a Gaussian,



Then,



Want to extremize this. So,



Plugging this back in, we get our energy,



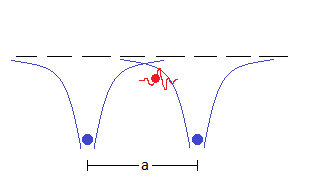
Okay. We could go on to estimate the first excited state. We just need to minimize H in the space of functions orthogonal to ψ0(x). We could simply guess that ψ1(x) = xψ0(x), times some normalization factor, in analogy with the HO states. Or more formally, we could set up and minimize the functional:



where α is a Lagrange multiplier that requires the overlap between ψ1 and ψ0 to be 0. But this would just give us the Schrodinger equation back again, which we don’t want to try to solve. So I guess, practically, we’d go with the former approach. Last comment: we could easily do the same for any power of x in the potential.

**Example. Double finite potential well**

Consider a molecule consisting of two identical atoms, separated by a distance a. Let their two potentials be V1 and V2 [they will have same functional form]. And consider an electron (red guy) somewhat bound to the two atoms.



Then the H for the electron will be:



Furthermore, say we know the eigenstates of either atom separately,



What is the approximate ground state wavefunction/energy of the combination? This problem is not really amenable to perturbation theory. We cannot treat V2 as a perturbation on V1 or vice versa, because V2 is just as large as V1. Or even if we could, somehow, we’d have to add up all terms in the series probably. The BWTIPT suffers the same problem. This is where the variational principle comes in handy. So if a = ∞, then the ground state would be degenerate; it could be either |1> or |2> (these are the ground states of the two separate H’s), or a linear combination of the two. As we bring the atoms together, we may surmise that the true ground state is approximately still some linear combination of the two individual ground states [we saw this was true when we solved the two-delta function-potential, and in general this is true – this is how energy bands form in crystal lattices]. So we can posit for the *normalized* ground state:



where c is some to-be-determined constant, and S = <1|2>. Then we form the expectation of H.



and minimize it to determine c. But first, let’s specialize to make things easier. We’ll presume Parity symmetry. And we’ll also have that P|1> = |2> and vice versa (P is the parity operator). So,



And note that H12 must be real, since H12\* = H21 = H12. Also,



Also,



So then,



Now differentiating to minimize,



So we have:



and the energies are:



Now



and,



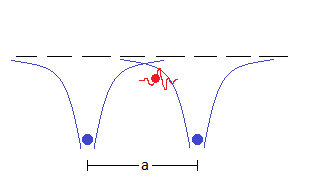
So then,



Note we got the first excited state as well. Further, since V is negative, the lowest energy state is the symmetric one, |ψ+>. In this state, the electron spends more time in between the two atoms than it does on either site per se´. This suggests that two atoms can lower their energy by sharing an electron – by bonding – than by any one of them ‘keeping’ the electron. Finally note that a so-called band gap opened up. The energy levels, formerly degenerate, have now split into two distinct levels, and the band gap is roughly 2<1|V1|2>. We’ll find this general phenomenon persists when we study periodic crystals.

**Example. Double finite well reprised…doing it that other way.**

Let’s reconsider the problem.



And H will be, for the electron:



We noted that both |1> and |2> are approximate eigenstates of H. And we said that the (better approximation to the) true ground state |ψ> lay in the subspace spanned by these two approximate eigenstates.



So one way to get |ψ> is to project H onto this subspace and solve it.



So we could act <1| and <2| respectively on both sides of this equation and get a matrix equation. But this will give us E’s in all four components of the matrix equation (because remember |1> isn’t orthogonal to |2>). So I’m going to do a Gram-Schmidt like process to simplify our equation beforehand. Let’s call the state within the subspace spanned by |1> and |2>, which is orthonormal to |1>, to be |1´>, and the one within the subspace orthonormal to |2> to be |2´>. These would be (unnormalized):



(where recall Parity makes S real) Now let’s act on both sides of our Schrodinger equation with <1´|,



and now with <2´|,



So altogether, recalling H12 = H21, and H11 = H22, thanks to Parity assumption,



which is,



Energies would follow from,



which is precisely what we got before!

**Example. Double finite well reprised again…doing it the ‘tight-binding’ way.**

Now, let’s suffice to say that if the overlap S = <1|2> between the wavefunctions is small *enough*, then we can quickly get a matrix equation by forming a projection operator for our subspace: I = |1><1| + |2><2|. So,



Then projecting against |1>, |2>, we’d have:



We argued before that on symmetry grounds, this is:



and employing our notation above, we can write this as:



So generically the eigenvalues are:



This result agrees with the above, to zeroth order in S. Now the eigenvectors would be:



Either one of the rows is now linearly depedendent on the other. So we can eliminate one (say bottom one) by linear operations. Whatever. So now we have, say:



where N is some normalization. And,



So then we’d have:



And so the new wavefunctions are just symmetric/antisymmetric combinations of the original ones. This also matches our previous result to O(1).